Introduction to Vector Space Models - Worksheet

Part One

1. Is the vector $\mathbf{x} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ in the $span \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$?

2. Is the vector $\mathbf{x} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ in the $span \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$?

3. Describe the span of one vector in $\mathbb{R}^3$.

4. Describe the span of two linearly *independent* vectors in $\mathbb{R}^3$.

5. Describe the span of two linearly *dependent* vectors in $\mathbb{R}^3$.

6. Compare the $span \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ to the $span \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\}$

7. What is the dimension of the $span \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\}$?

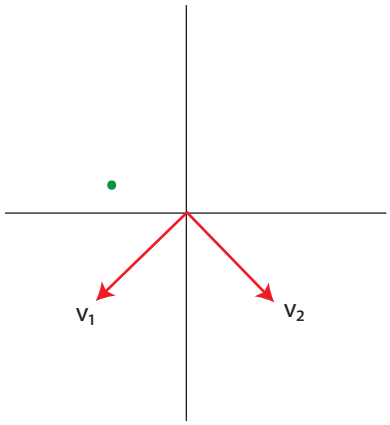
8. What is the definition of the **dimension** of a subspace?

9. How would you describe a hyperplane?

Part Two

1. What are the coordinates of the vector $\mathbf{x} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ in the basis $\left\{ \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$? Draw a picture to make sure your answer lines up with intuition.

2. In the following picture what would be the signs (+/-) of the coordinates of the green point in the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$? Pick another point at random and answer the same question for that point.



Part Three

1. Interpret the following Nonnegative Factor Output for a small collection of text documents, answering the following questions:

   a. What meaning (theme/topic) would you give to each of the three factors?

   b. What is the dominant factor (theme/topic) for each document?

   c. What is the loading of the word *baseball* on Factor 2?

   d. What is the coordinate/score of document 5 along Factor 3?

$$TermDocMatrix \approx \begin{array}{c} \\ \text{“baseball”} \\ \text{“pitcher”} \\ \text{“mound”} \\ \text{“player”} \\ \text{“coach”} \\ \text{“soccer”} \\ \text{“world”} \\ \text{“fifa”} \\ \text{“cup”} \\ \text{“canada”} \\ \text{“womens”} \\ \text{“USA”} \\ \text{“olympics”} \\ \text{“medal”} \\ \text{“gold”} \\ \text{“phelps”} \end{array} \begin{array}{c} \begin{array}{ccc} \textit{Factor}1 & \textit{Factor}2 & \textit{Factor}3 \end{array} \\ \begin{pmatrix} 1.9 & 0 & 0 \\ 2.6 & 0 & 0.1 \\ 1.1 & 0.0 & 0 \\ 1.5 & 0.1 & 0 \\ 1.3 & 0.8 & 0.8 \\ 0 & 2.2 & 0 \\ 0.1 & 1.7 & 0.5 \\ 0 & 2.3 & 0 \\ 0 & 1.6 & 0.1 \\ 0.2 & 1.9 & 0.5 \\ 0 & 1.8 & 0.7 \\ 0.1 & 2.0 & 2.3 \\ 0 & 0.2 & 2.8 \\ 0 & 0.1 & 2.2 \\ 0 & 0 & 1.8 \\ 0 & 0 & 1.6 \end{pmatrix} \end{array} \begin{array}{c} \begin{array}{ccccc} \textit{doc}1 & \textit{doc}2 & \textit{doc}3 & \textit{doc}4 & \textit{doc}5 \end{array} \\ \begin{pmatrix} 3.2 & 2.7 & 0 & 0.2 & 0.1 \\ 0.1 & 0.1 & 2.5 & 2.1 & 0.3 \\ 0.2 & 0 & 0.2 & 0.1 & 2.9 \end{pmatrix} \end{array}$$

List of Key Words.

linear combination geometrically
linear (in)dependence geometrically
vector span
subspace
dimension of subspace
hyperplane
basis vectors
coordinates in different bases
(generic) factor analysis
loadings
scores/coordinates